语言学研究

人类语言习得的亲知还原模式
——从 ChatGPT 的言知还原模式说起

陈保亚[1]　陈　樾[2]

(1. 北京大学 中国语言学研究中心, 北京 100871;2. 佐治亚大学 数学系,美国)

摘　要: 尽管语言人工智能的大语言模型 ChatGPT 取得了比较大的进展,哲学上的图灵和塞尔之争仍然在继续。不过 ChatGPT 能够生成符合语法的崭新的句子,一定还原出了语言单位(tokens)和规则,解决了长期以来人工智能中自然语言理解的难题,这是一个重要的转折。ChatGPT 的学习模型依赖强大的运算能力和计算机的海量存储能力,这两种能力可以合称为强储算能力。相比之下,人脑只具有弱储算能力。正是因为弱储算能力的限制,人脑语言学习不可能完全走 ChatGPT 的语言学习模式。人脑是在基于经验的亲知活动中还原出有限的单位和规则,从而生成崭新的句子。ChatGPT 目前采用的是言知学习模式,而不是基于经验的亲知学习模式,将来的大语言模型可能扩展出亲知学习模式,真正模拟人类获得亲知还原模式。那个时候或许可以说机器人真正理解了自然语言,哲学上的图灵和塞尔之争或许可能得到解决。

关键词: 人工智能;图灵测试;中国房间;自然语言理解;思维

中图分类号:H08　　**文献标识码:**A　　**文章编号:**1000-5919(2024)02-0167-08

一、缘起:图灵和塞尔之争

ChatGPT 在自然语言理解方面的强大功能引起了人们广泛关注。① 其中一个引起人们关注的重要问题是 ChatGPT 是否能够像人一样思维。图灵(Turing)在《计算机器与智能》中提出了模仿游戏(imitation game),后人称为“图灵测试(Turing test)”。基本思想是,在测试者 C、被试机器 A 和被试人 B 都相互隔绝的状态下,测试者 C 提出各种问题,被试机器 A 和被试人 B 回答问题,测试者 C 如果不能区分 A 和 B 谁是机器,图灵就认为机器 A 能够思维。②

图灵测试本身还不够明确,因为很多问题人不能回答,机器反而能够回答。比如只要提问 6 开平方是多少,能够给出答案的应该是机器人,因此要区分人和机器其实是很容易的。不过图灵的基本思想是清楚的,只要机器能够完成人能完成的大部分工作,机器就可以说能够思维了。就一般问题而言,基于 GPT-3.5 的 ChatGPT,基本上能够完成图灵测试,可以说已经符合图灵所说的思维条件。这是否可以说 ChatGPT 能够思维?

20 世纪 80 年代,塞尔在《心智、大脑和程序》中针对图灵测试提出了中国房间(Chinese room)的问题。概括地说,假设把图灵测试中的机器人换成一个不懂中文的英语人,只要借助英语写的关于汉语的手册等工具,该英语人就可以根据汉语句子的提问用汉语句子回答。但这并不等于机器懂汉语,也并不

收稿日期:2023-12-01

作者简介:陈保亚,北京大学中国语言学研究中心研究员、北京大学中文系教授。

基金项目:国家社科基金重大项目“我国民族音乐文化与语言数据集成及共演化研究”(22&ZD218)

① 冯志伟、张灯柯、饶高琦:《从图灵测试到 ChatGPT——人机对话的里程碑及启示》,《语言战略研究》2023 年第 8 期,第 20—24 页。

② Turing, Alan,“Computing Machinery and Intelligence”, *Mind*, LIX(236),October, 1950, p. 433-460.

能证明机器懂汉语。① 塞尔的中国房间被称为思想实验。在中国房间实验中,英语人并不懂汉语,不理解汉语,却可以输出汉语句子。塞尔想表达的观点是,机器能输出句子并不等于机器理解了句子。塞尔的思想实验中英语人是理解英语的,汉语句子的输出有语言人在操作,尽管不是汉语人。现在的问题是,ChatGPT 完成汉语和英语翻译,生成新的句子和文本,并不需要语言人的操作,这算不算 ChatGPT 能理解语言?

二、ChatGPT 语言学方法:基于言知的分布理论

ChatGPT 目前还有很多能力无法和人相比,比如数学证明能力。ChatGPT 的有些对话文本可能让人感到莫名其妙。但可以肯定的是,ChatGPT 说出的每个新句子都是符合语法规则和语义规则的,它不会说出“见面老师”这样的违反语法的表述,也不会说出“喝牛排”这样的违反论元语义规则的表述。这一点证明了 ChatGPT 肯定还原出了语言单位和规则,否则 ChatGPT 无法生成符合语法语义规则的崭新句子。这不能不说是自然语言处理上的重要转折,在此之前,让机器像人一样说出正确的句子,一直是计算语言学的难点问题。

ChatGPT 是一种大语言模型(large language model),有两个重要的基础,一个是人工神经网络(artificial neural networks)数学模型,一个是大数据。人工神经网络又简称神经网络,其基本原理就是模仿人类大脑神经网络进行非线性回归计算,自动建立预测模型。这种神经网络有很多层次,这些层次隐藏很深,所以其自动建模的过程也称为深度学习(deep learning)。ChatGPT 的神经网络由数十个网络层构成,每个层都是一个转换器(transformer)。转换器至关重要,是瓦斯瓦尼(Vaswani)等在人工智能奠基性论文《关注就是你所需要的》(Attention is all you need)中提出来的②,有效解决了自然语言特征的提取技术。ChatGPT 中的 T,就是 transformer 第一个字母的缩写。和此前机器人比较起来,人工神经网络模仿人的大脑机制,更贴近人脑的语言学习模式。瓦斯瓦尼等论文中的“关注”(Attention)就是用数学方式描写一个词和另一个词的密切关系。ChatGPT 还包括了 RLHF(Reinforcement Learning Human Feedback),即人类反馈强化学习,这样可以不断调整自己,更接近人类的行为模式。

基于人工神经网络,ChatGPT 能够自己读取网络上包括维基百科等在内的海量文本,从中模拟出语言生成模型。ChatGPT 在数学计算和数据存储方面,远远超过了真人的大脑,正是因为有这两个超强的能力,自然语言中的词语可以真正被自动标注,实现高维向量化,形成复杂网络关联,人工神经网络可以进行大规模运算得到最佳输出。2013 年 Google 公布的 Word2vec 模型,主要就是在讲词的向量化。该模型的词向量已达 600 多个,即每个词都由 600 多个参数构成,每个参数都表达了一种词和其他词的相似关系。该模型可看成大语言模型的范式。通俗地看,词的向量化有些像语言学和计算语言学中的特征标注,比如“哥哥”和“姐姐”都可以标注“同辈亲属、年长”的特征,即两个向量,“哥哥”和“姐姐”于是有了相关性,在符合“同辈亲属、年长”的语境下,这些词就可能共现。所不同的是,语言学和计算语言学中的特征标注是人工的,“同辈亲属、年长”只是两个特征,自然语言中和语言组合规则相关的语言特征远远不止两个,词与词之间的聚合关系和组合关系也不是仅仅靠少量特征就可以描述。要进行大量的特征标注,难度就会剧增。而且标注词的哪些特征才能反映句子生成规则,人工标注并不是很清楚,

---

① Searle, John, “Minds, Brains and Programs”, *Behavioral and Brain Sciences*, 1980, 3 (3), pp. 417-457.

② Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin, “Attention Is All You Need”, in *Neural Information Processing Systems*, Cambridge, USA: MIT Press, 2017, pp. 6000-6010.

必须不断手工调试,每次调试都是巨大的工作量。基于 Chat-3.5 的 ChatGPT 是自动识别、自动调试、自动反馈和自动输出的最佳模拟装置,解决了海量计算难题。比如,瓦斯瓦尼等提出 Transformer 的论文里,反映词项分布特征数量的向量维度已经高达 512,到 GPT-3,向量维度已经达到 12288,神经网络有 96 层。所包含的信息已经相当大。GPT-4 尽管没有公布向量维度,但肯定不弱于 GPT-3。更重要的是,ChatGPT 向量化的大规模扩充、储存和运算,仍然是通过神经网络自动实现的,这就为机器在海量文本中自动建立词的特征标注或向量化提供了可能。也正是通过词的向量化,ChatGPT 可以进一步获取词的分布概率,建立词的组合模型,完成生成新句子的工作。

尽管 ChatGPT 的人机对话在内容上还存在不足,但从语言理论的角度看,ChatGPT 能生成崭新的符合语法的句子和文本,这在自然语言理解上是一个非常值得重视的进展,因此可以肯定,ChatGPT 能在现存的文本语料中还原出单位和规则,生成崭新的句子。如果 ChatGPT 只是凭借能处理海量数据的能力死记硬背语言句子,不可能生成崭新的句子和文本。由于人工神经网络内部是一个黑箱,ChatGPT 是如何自动建立向量空间,如何建立语言生成模型的,人们并不清楚。但有一点很明确,ChatGPT 并不跟经验打交道,而是直接面对海量文本。海量文本可以提供每个词足够详细的分布信息,ChatGPT 可以而且只能依靠词的分布通过人工神经网络还原单位和规则,建立语言生成模型。从数学方法看,人工神经网络本质上就是一种非线性回归的算法,只要输入材料足够的丰富,这种算法便能够模拟出材料背后的规律,形成自动建模。ChatGPT 把词的分布转化成数学中的向量,借助海量的语言文本,最终模拟出语言规则,并用这些规则生成符合语法的崭新句子和文本。人工神经网络的方法原则从源头上看就是数学中勒让德(Legendre)和高斯(Gauss)的回归理论,只不过勒让德和高斯的回归是线性回归,后来数学家发展出了非线性回归,但基本原则是相同的,即如何从复杂的要素分布中模拟出数学模型,然后预测未知的要素分布。这是一种分布理论。

从语言理论的角度看,ChatGPT 的语言学方法也是分布理论,这种思路正是海里斯(Harris)分布理论的思路。更早的时候,布龙菲尔德(Bloomfield)的语言行为主义理论,也是把词的意义看成词的用法(分布)。① 海里斯认为,只要充分描写每个语素(morpheme)的分布规则,就可以得到该语言的规则。② 由于语言中语素的分布异常复杂,几乎每个语素都有自己不同的分布,所以海里斯当时不可能充分描写一个语言全部语素的详细分布,而只能以举例的方式概括出分布理论。语言学家也难以靠手工穷尽语素的分布。ChatGPT 使这种大规模分布运算成为可能,无论是语素,由语素构成的词,还是其他语言单位(tokens),大规模运算都是可能的。

海里斯的分布理论是一种独立于经验的纯形式分析。既然 ChatGPT 能够独立于经验获得分布规则,这也验证了形式主义语法理论,即和语义相对的语法规则可以独立于经验,这正是海里斯的学生乔姆斯基(Chomsky)的基本思想。③ 自此我们可以形成一种关于符号系统的认识。自数学、逻辑学中公理化系统形成后,人们已经认识到数学是一种纯形式公理系统,不需要语义解释,通俗地说就是不需要

---

① Bloomfield, L.,“A Set of Postulates for the Science of Language”, *Language*, Vol. 2, 1926, pp. 153-164; Bloomfield, L., *Language*, New York: Henry Holt, 1933.

② Harris, Z. S.,“From Morpheme to Utterance”, *Language*, *Vol.* 22, 1946, pp. 166-183; Harris, Z. S., *Methods in Structural Linguistics*, Chicago: University of Chicago Press, 1951.

③ Chomsky, N., *The Logical Structure of Linguistic Theory*, New York: Plenum, 1975, p. 573. Chomsky, N.,“Three Models for the Description of Language”, I. R. E. *Transactions on Information Theory*, Vol. IT-2, 1956, pp. 113-124; Proceedings of the Symosium on Information Theory, Sept, 1956; Chomsky, N., *Syntactic Structures*, the Hague: Mouton, 1957, pp. 1-15.

经验的支持,只要公理相互无矛盾即可,这样数学从本质上是一种纯形式符号系统的思想已经确定。至于数学和实际应用的关系,取决于实际的需求,比如欧氏空间需要欧氏几何,相对论需要非欧几何,等等。ChatGPT 在语言生成上的成功是独立于经验的,这也证明自然语言中存在独立于经验的形式系统。至于这种形式系统如何应用,则需要和经验联系。ChatGPT 所做的人工训练和监督,属于和人类价值观、道德水平对齐(alignment),这部分就好比儿童学会了语言,仍然要对他们进行道德、价值观和法律等教育。这些人工训练和监督并不属于语言能力的训练。

三、人类语言习得:亲知还原能力

回到图灵和塞尔之争:ChatGPT 是否能思维,是否理解了自然语言?

回答这个问题取决于我们怎么定义思维和语言,但有一点值得注意,尽管 ChatGPT 从海量文本中获取了一种语言生成模型,但 ChatGPT 的自动学习方式和人类的语言学习方式并不相同。

前面提到,人脑无论在存储能力和运算能力上都远远不如 ChatGPT,只具备弱储算能力。正是因为弱储算能力的限制,人脑语言学习不可能完全走 ChatGPT 的语言学习模式。人脑同样可以像 ChatGPT 一样生成崭新的句子和文本,不过人脑语言生成能力是从有限的并且较少的规则和单位生成崭新的句子和文本。要完成从有限到无限,人脑需要在基于经验的亲知活动中还原出有限的单位和规则,从而生成崭新的句子和文本。人脑还原单位和规则所用到的文本数量和 ChatGPT 所用到的文本数量要远远小得多。基于亲知还原出来的单位和规则是根植于经验的亲知语言。相比之下,ChatGPT 目前采用的是言知学习模式,而不是基于经验的亲知学习模式。ChatGPT 通过多维度向量进行大规模计算,也能从浩瀚的语料中总结出关于外部世界的很多信息,但这些信息是靠言知而不是亲知获得的。从储算能力看,ChatGPT 可以还原出形式单位(tokens),但能否还原出基于经验的有意义的单位还不得而知。将来 ChatGPT 可能扩展出亲知学习模式,比如在嗅觉、触觉、痛感、忧伤感、愉快感以及通感等方面取得进展,从而模拟人类获得最佳还原模型,那个时候或许可以说机器人真正有了思维,真正理解了语言。

为什么 ChatGPT 需要超大数据的学习才能学会怎样生成崭新的句子,而人类只用了有限的数据就能学会理解和生成崭新的句子? 三岁小孩基本上已经掌握了母语,他们所接触到的句子数量是相当有限的。显然,由于人类的弱储算能力,人类只能在一定数量的句子中(包括独词句)还原规则和单位,再依据这些有限的单位和有限的规则,生成崭新的句子。具体地说,儿童通过生活游戏获得一些词和句子的用法,这只是言语习得的第一步,第二步是单位和规则的还原过程。还原过程是类推。比如儿童可以学会以下词组:

布鞋、草鞋、皮鞋

金表、铜表、银表

儿童会以类推的方式还原出这里的单位,并生成新的组合:

金鞋、铜鞋、银鞋

这里的类推结果是建立起一个基于共相的模型“X 鞋”,其中 X 表示质料。类推的本质是用已经有的知识言说新的知识。这里的知识是和经验世界相关的,即“质料”知识的形成是人类在和经验世界打交道时形成的亲知。如果没有基于经验的类推,自然语言的还原过程和生成过程都不能实现。概括地说,人类知识不仅有言知,而且还有亲知,并且言知是建立在亲知基础上的。

类推的核心在于是否能够周遍类推。有些类推是不周遍的,有些类推是周遍的①:

平行周遍模式:白纸、白墙、白鞋、白毛、白车……#白菜、#白金、#白铁…… $ 白煤……

平行不周遍模式:腕儿、腿儿、桌儿、门儿、本儿、嘴儿、肝儿、肠儿、#心儿,#眼儿…… * 笔儿、* 墨儿、* 脚儿、* 手儿、* 掌儿、* 指儿、* 鼻儿、* 牙儿、* 胃儿、* 肾儿……

平行周遍模式中遇到的反例通常是可解释的。一种是转义,比如上面标记有符号#的“白菜、白金、白铁”。另一种情况是经验知识中还不存在的组合,比如上面标有符号 $ 的“白煤”。除了这些反例,“白X”是可周遍类推的模式,这种模式可用来创新。崭新句子的生成就是基于这种周遍类推。

上面平行不周遍模式“X 儿”是不作为规则来生成新实例的。比如“腿儿”和“脚儿”在组合关系上是平行的,“腿儿”可以说,“脚儿”却不可以说。只有平行不周遍模式的实例才需要记忆,平行周遍模式不需要记忆,这就是还原学习的充分性。这两种模式的实例计算机可以完全入库,不需要还原出更小的单位,因为计算机有足够的强储算能力。这是计算机自然语言处理不同于人脑处理语言的重要区别。

ChatGPT 能否像人脑一样进行语言和规则的还原? 目前看来是有困难的。人脑能够进行充分的规则和单位的还原,依赖了亲知,这可能是人类能够在弱储算能力下也能充分还原单位和规则的关键。ChatGPT 等所依赖的人工神经网络目前还不具有亲知,只有言知。比如百度推出的大语言模型文心一言,和 ChatGPT 一样以人工神经网络为基础,其定义“质料”用到了“物质”,定义“物质”用到了“实体”,定义“实体”用到了“实物”,定义“实物”用到了“物体”,定义“物体”用到了“实体”,最终形成循环解释。这些词之所以用来相互定义,是因为人工神经网络在大规模的句子中发现“质料、物质、实体、实物、物体”等词经常出现在句子的相同位置,因此被置于相同的向量空间中,具有相似性。相似性就可以提取相似特征。根据词的分布提取词的相似性,这就是大语言模型的基本工作原理。而日常生活中的人对“质料”的理解,对“物质、实体、实物、物体”的理解,是通过触觉、视觉等在经验中获得的感知结果。类比地说,计算机可以通过“芒果”“香蕉”“菠萝蜜”“榴莲”等词在大数据中的分布,自动建立起“芒果”“香蕉”“菠萝蜜”“榴莲”的相似关系,提取“水果”的特征,但仍然不知道这些水果真正的味道是什么。

用言知的方式还原词的意义必然陷入循环,比如“知道”,《现代汉语词典》(第 7 版)的解释是循环的:

知道:对于事物和道理有认识;懂得(p1678)

懂得:知道(意义、做法等)(P312)

认识:1. 能够确定某一人或事物是这个人或事物而不是别的。2. 通过实践了解、掌握客观事物(P1102)

了解:知道得清楚(P820)

以上“知道、懂得、认识、了解”都是在循环定义,循环解释。显然,人类不是通过定义、解释这样一些言知学习模式而获得“知道”的意义,而是通过亲知过程中的语言使用。使用先于定义,亲知先于言知。自然语言是最初始的基于亲知的元语言,给词下定义最终是一个循环过程。

在命题逻辑中有如下定义:

¬(否定符号,¬ p 表示“非 p”),

这是用自然语言“否、非”定义符号¬ 。但自然语言“否”“非”和“不”在《现代汉语词典》又是循环定义:

---

① 陈保亚:《论平行周遍原则与规则语素组的判定》,《中国语文》2006 年第 2 期,第 99—108 页。

否:表示不同意

非:不、不是、不合于

不:用在动词、形容词和其他副词前面表示否定

逻辑学中最重要的判断词,《现代汉语词典》中的定义也是循环的:

是:对,正确

对:相合,正确,正常

正确:符合事实

符合:相合

人类自然语言以及在自然语言基础上建立起来的符号系统,从语言层阶的角度看都存在着循环解释。维特根斯坦说:“我如果不能给‘植物’下个定义我就不知道自己在说什么吗?”①这也是承认使用先于定义,亲知先于言知。

我们再回到图灵和塞尔关于“思维(thinking)”和“理解(understanding)”这个问题上来,其实图灵的功能派和塞尔的解释派都没有对“思维、理解”做严格定义,我们需要严格区分两种“思维、理解”。图灵以及人工智能所说的机器人能够思维,能够理解语言,这种思维和理解是指词与词之间的网络关系,塞尔所说的思维和理解,是指基于经验的思维,是指对词语背后经验世界的理解。

我们可以把人类通过亲知学会的语言称为亲知语言,而把机器人通过言知学会的语言称为言知语言。自然语言是通过初始词项及其不断扩充而获得的符号系统,这种扩充既包括隐喻用法或类比用法,也包括定义和解释,但初始语言系统本身不可能通过定义和解释获得,只能通过亲知获得。机器人将来能否实现亲知的学习模式,获得亲知语言,是自然语言理解的关键。目前机器人对真人的图像识别和声音的识别,是一种亲知的开始。但至少在目前情况下,要达到人类亲知世界的程度还是有相当难度的,因为电脑和人脑的构造并不一样。人脑是一种有机构造,后面有复杂的生物结构,还连带着身体各部分的复杂生物结构,人类正是通过这些复杂的生物结构和经验世界打交道而获取亲知语言。人的复杂生物结构是在长期的进化过程中逐渐形成的,具有复杂的感知世界的亲知系统,这是目前的机器人还不具备的。ChatGPT 的飞跃的一个关键是大数据计算,这些大数据来自于研究机构收集的数据、网上的数据等,包括网上维基百科的数据等。网上的数据是良莠不齐的,必然影响 GPT 的质量。微软的研究人员古纳塞克拉(Gunasekar)等在微软预印本服务器 arXix 上发表了文章《文本是你所需要的全部》(Textbooks Are All You Need),强调要提高数据质量。在 2023 年 1 月公布的 phi-1 模型中,提高数据质量后模型的性能明显提高。GPT 5 也将遵循提高数据质量的思路,但是,文本质量再高,也是基于言知的数据,模型终究还是基于言知的模型,和人类语言学习方式不一样。一种可能性的发展趋势是物理机器人向生物机器人发展,机器人逐渐发展出亲知能力。现在的计算机对外部世界的音像识别可以看成是亲知能力的先兆。

四、结语:语言行为与语言认识

基于大语言模型的机器人可以在没有亲知的情况下,根据已有文本能够获得句子生成能力,生成崭新的句子,并在这种言知语言能力的基础上完成大量信息处理、推理工作和创作活动。因此,我们关于语言能力、语言知识和思维能力的理论需要调整。至于图灵和塞尔关于机器是否有思维和理解能力,取决于我们如何定义“思维”和“理解”。乔姆斯基认为 ChatGPT 并没有告诉我们任何语言知识,如何理解

① Wittgenstein, L. *Philosophical Investigations*, 70, Oxford: Basil Blackwell, 1953, p. 33.

这一点,则取决于如何定义"语言知识"。绕开这些争论,有一个基本问题是清楚的,ChatGPT 的言知储算模型必须要依靠强储算能力才能还原语言单位和规则。人类学习语言,不可能用 ChatGPT 的方法。亲知储算模型只需要弱储算能力即可还原语言单位和规则,这是人类学习语言的特点,这背后真正的机制语言学家、人工智能专家都还没有研究清楚。也许正是这种亲知储算模型,使得人类具有另一些能力,其中最重要的就是数学证明能力和在怀疑、反思、领悟的基础上建构相对论这样一类理论的能力。这是人工神经网络目前还未实现的,今后是否能实现需要进一步研究。

即使将来机器人通过深度学习发展出了亲知能力和亲知语言,也并不能代表人类对语言机制有了充分的认识,这只是语言行为的实现。ChatGPT 涌现出的一些高度复杂的行为模式,人工智能专家目前也不能完全认识这些行为的机制。认识人类语言的运转机制永远是语言科学研究的目标,这种认识可能对推进机器人学会亲知语言提供理论支持,也为我们限制机器人产生破坏行为提供了借鉴。一旦机器人能够完全像人类那样以亲知还原的方式学习语言,加上机器人自己的强储算能力,机器人的语言能力和思维能力是惊人的。而机器人学习语言的过程,也为人类认识语言运转机制提供了更多的窗口。比如,ChatGPT 学习语言并没有使用大量的语法术语和跟语法体系相关的概念,都提示我们语法研究更应该重视共时规则和历时规则的研究,重视语言运转机制的研究,而不是一味建构抽象复杂的语法体系。语言学不仅要研究人类学习语言的机制,还需要研究机器人学习语言的机制。

我们现在的语言学,是基于人类学习语言的语言学,机器人学习语言的机制也应该纳入研究,因此我们需要有一种基于既研究人类语言,也研究机器人语言的广义语言学,这样我们关于语言学习的机制才更充分。

ChatGPT 还存在很多不足,但它在多方面的成功是不容忽视的,这些成功显示了人工神经网络在自然语言理解中的重要性。自然语言理解早期经历了规则模型和概率模型,都遇到了很多困难。规则模型的困难不在于规则本身不重要,而在于人工寻找和建立规则的困难。这里的教训是,会说语言的人不是很容易找出语言背后规则的,就像会消化的人并不容易找到消化的规律。其实 ChatGPT 也是在寻找规则,不过是在大运算和大数据中寻找规则,这从一个侧面反映了寻找规则的难度。概率模型也是很重要的,人工神经网络的计算就是基于概率模型,只是早期的概率模型也遇到了对海量文本进行大规模运算的困难。规则模型和概率模型都是有价值的模型,问题是怎么去自动实现,人工神经网络做了自动实现的工作。

古人类学家、考古学家、古遗传学家等都认为符号系统的产生是人类进化的一个重要特征,但并不认为是一个最为关键的特征,并没有拿符号来作为一个区别性特征,这令人遗憾。现有的实验显示,动物在很多方面都具有超过人类的能力,比如海豚用声波定位的能力,黑猩猩的记忆能力等等,但是这些动物都未能发展出高度发达的人类社会。我们认为人类进化的一个根本转折是掌握了语言。有了语言这样一种符号系统,经验才可能有序化,人类才可能做出未来计划,才可能分享创新,才可能积累知识。人类进化最为关键的环节是符号系统的产生,即能产符号系统的产生。① 自然语言就是一种符号系统。卡西尔曾经把人定义为符号动物②,现在看来是有道理的。OpenAI 高度关注人工智能中的自然语言理解,其大语言模型 ChatGPT 在人工智能中实现了自然语言的生成,这是一个重要的转折。ChatGPT 的出现对机器人发展的重要性犹如自然语言符号系统的出现对人类发展的重要性,从此机器人可以和人类

---

① 陈保亚:《语言与音乐:认识人类听觉符号的两座基石——从人类进化的符号能力说起》,《贵州民族大学学报(哲学社会科学版)》2023 年第 4 期,第 5—13 页。

② 卡西尔(Ernst Cassirer):《人论》,甘阳译,上海:上海译文出版社 1985 年版。

进行自然语言对话,直接阅读、继承和使用人类用自然语言记录的浩瀚文本知识。语言是思维和交际的最重要的工具,想绕开语言来实现人工智能,就好比绕开语言来谈人类起源,是方向性的错误。可以说,语言人的出现是人类进化的转折点,语言机器人的出现是机器人历史的转折点,这一转折引导我们开始思考更多的语言理论问题。

The Model of Familiarity Reduction in Human Language Acquisition

—— From the Viewpoint of ChatGPT's Model of Verbal Reduction

Chen Baoya[1], Chen Yue[2]

(1. *Center for Chinese Linguistics*, *Department of Chinese Language and Literature*, *Peking University*, *Beijing* 100871 *China*;
2. *Department of Mathematics*, *the University of Georgia*, *USA*)

Abstract: Despite the significant progress made in ChatGPT, a big language model of artificial intelligence, the philosophical debate between Turing (图灵) and Searle (塞尔) is still going on. However, ChatGPT is able to generate new sentences that conform to grammar, and has certainly reduced language units (tokens) and rules, thus solving the long-standing problems of natural language understanding in artificial intelligence. This is an important turning point. The learning model of ChatGPT relies on strong computing power and the massive storage capacity of computers, which can be collectively referred to as strong power of storaging and computing. In contrast, human brain has only weak of storaging and computing power reduce. It is precisely because of the limitations of its weak power of storaging and computing that human brain language learning cannot completely follow the language learning model of ChatGPT. Human brain reduce limited units and rules through activities of familiarity based on experience, thereby generating new sentences. ChatGPT currently adopts a text-based learning model, rather than an experience-based learning model of familiarity. The future bigger language models may expand to include a learning model of familiarity, truly simulating the model of familiarity reduction in human language acquisition. Until then, it may be said that robots can truly understand natural language, and the philosophical dispute between Turing and Searle may have been resolved.

Key words: artificial intelligence; Turing test; Chinese room; natural language understanding; thinking

(责任编辑　管　琴)